



PATTERNS OF ASSOCIATION – QUALITY OF ENGLISH SPOKEN BY PEOPLE WHO SPEAK SPANISH IN THEIR HOMES

TEACHER VERSION

Subject Level:

Middle School Math

Grade Level:

8

Approx. Time Required:

45-60 minutes

Learning Objectives:

- Students will be able to create a scatter plot from a data table.
- Students will be able to interpret a scatter plot to determine associations between two quantitative variables.

Activity Description

Students will create a scatter plot using data from the U.S. Census Bureau’s American Community Survey (ACS), in which residents who primarily speak Spanish at home were asked to self-report the quality of their spoken English. Students will study associations between the years the survey data were reported (2015–2019) and the ratings for speaking English (“very well” and “less than very well”).

Suggested Grade Level:

8

Approximate Time Required:

45–60 minutes

Learning Objectives:

- Students will be able to create a scatter plot from a data table.
 - Students will be able to interpret a scatter plot to determine associations between two quantitative variables.
-

Topics:

- Coordinates of a graph
- Positive and negative associations
- Scatter plots

Skills Taught:

- Interpreting graphs in context
 - Determining strength of association
 - Understanding associations between quantitative variables
-

Materials Required

- The student version of this activity, 9 pages
- Colored pencils

Activity Item

The following item is part of this activity and appears within this teacher version.

- Item 1: Quality of English Spoken in Primarily Spanish-Speaking Homes in the United States, 2015–2019

For more information to help you introduce your students to the Census Bureau, read [*“Census Bureau 101 for Students.”*](#) This information sheet can be printed and passed out to your students as well.

Standards Addressed

See charts below. For more information, read

[*“Overview of Education Standards and Guidelines Addressed in Statistics in Schools Activities.”*](#)

Common Core State Standards for Mathematics

Standard	Domain	Cluster
CCSS.MATH.CONTENT.8.SP.A.1 Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.	8 SP – Statistics & Probability	Investigate patterns of association in bivariate data.

Common Core State Standards for Mathematical Practice

Standard

CCSS.MATH.PRACTICE.MP4. Model with mathematics.

Students will use words to describe the associations they observe from the scatter plots in context.

CCSS.MATH.PRACTICE.MP7. Look for and make use of structure.

Students will observe positive and negative trends and translate them into words.

National Council of Teachers of Mathematics' Principles and Standards for School Mathematics

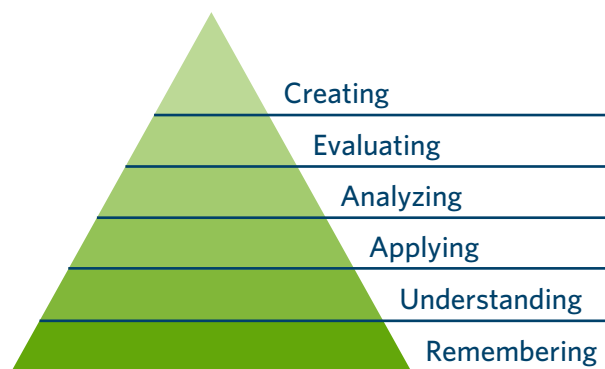
Content Standard	Students should be able to:	Expectation for Grade Band
Data Analysis and Probability	Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them.	Select, create, and use appropriate graphical representations of data, including histograms, box plots, and scatterplots.
Data Analysis and Probability	Select and use appropriate statistical methods to analyze data.	Discuss and understand the correspondence between data sets and their graphical representations, especially histograms, stem-and-leaf plots, box plots, and scatterplots.

Guidelines for Assessment and Instruction in Statistics Education

GAISE	Level A	Level B	Level C
Formulate Questions	X		
Collect Data			
Analyze Data		X	
Interpret Results		X	

Bloom's Taxonomy

Students will **apply** graphing skills to create a scatter plot and will **analyze** the direction, form, and strength of the associations.



Teacher Notes

Before the Activity

Students must understand the following key terms:

- **Bivariate data** – pairs of linked numerical observations. Example: a list of heights and weights for each player on a football team
- **Categorical variable** – a variable that is not numerical, such as a name or label, that places an object into one of several groups or categories (e.g., the color of a ball or the breed of a dog)
- **Explanatory variable** – the “independent” variable, which helps explain the changes in the response variable
- **Form of association** – the shape that data make moving from left to right on a scatter plot (e.g., linear, parabolic, quadratic)
- **Line of best fit** – a straight line drawn through the center of a group of data points on a scatter plot, showing how closely the two variables on the scatter plot are associated
- **Negative direction of association** – an association in which one variable increases as the other decreases; when a line of best fit represents a negative association, the line has a negative slope (e.g., for time a musician spends practicing a piece and mistakes made during performing; the more time spent practicing, the fewer mistakes that musician will likely make)
- **Outlier** – a data point that is well outside of the expected range of values or does not follow the overall pattern of the other data points
- **Positive direction of association** – an association in which one variable increases as the other increases; when a line of best fit represents a positive association, the line has a positive slope (e.g., for children’s age and height; as children get older, they usually get taller)
- **Quantitative variable** – a variable that is numerical, meaning that it represents a measurable quantity (e.g., the population size of a city)
- **Response variable** – the “dependent” variable, which shows an outcome
- **Scatter plot** – a graph in the coordinate plane that displays a set of bivariate data and can be used to determine how two variables are associated (e.g., to show associations between the heights and weights of a group of people)
- **Strength of association** – a measure of how tightly points are clustered

Students should have a basic understanding of the following concepts:

- Creating scatter plots
- Describing direction, form, and strength of associations on scatter plots

During the Activity

As students work, teachers will ask essential questions:

- Why are we using different colors in the legend?
- What patterns do you see in the data?
- Do you think the paths ever crossed in earlier years?
- Do you think the paths will cross in future years?
- What is the direction of association for the “very well” path? Could you explain why?
- What is the direction of association for the “less than very well” path? Could you explain why?
- Could you explain why one line has a positive direction of association and the other has a negative direction of association?

Once all students are finished with question 9, instruct them to complete question 10 with a partner.

After the Activity

Teachers should ask students to reflect on what they learned.

Extension Ideas

- Teachers could direct students to create a stacked bar graph representing the proportion of “very well” and “less than very well” percentages for each year.
- Teachers could ask students to produce their own data and analyze associations. Some examples are:
 - Handspan and height
 - Handwriting speeds with students’ dominant and nondominant hands
 - Number of jumping jacks completed and pulse (measured for 10 seconds afterward)

Student Activity

Click [here](#) to download a printable version for students.

Activity Item

The following item is part of this activity and appears within this student version.

- Item 1: Quality of English Spoken in Primarily Spanish-Speaking Homes in the United States, 2015–2019

Student Learning Objectives

- I will be able to create a scatter plot from a data table.
- I will be able to interpret a scatter plot to determine associations between two quantitative variables.

The American Community Survey (ACS) is conducted monthly by the U.S. Census Bureau and is designed to show how communities are changing. Through asking questions of a sample of the population, it produces national data on more than 35 categories of information, such as education, income, housing, and employment. In this activity, you will create a scatter plot to examine the trends in the quality of English spoken by U.S. residents who primarily speak Spanish at home. These data were self-reported on the ACS questionnaire from 2015 through 2019. You will measure two variables: the percentage of Spanish-speaking residents who reported speaking English “very well” and the percentage of Spanish-speaking residents who reported speaking English “less than very well.”

In creating your scatter plot, think about the following questions: When one variable increases, does the other also increase? How quickly or slowly does one variable change as the other changes?

Part 1 – Create a Scatter Plot

Item 1, below, shows the percentages of the U.S. population speaking Spanish at home who reported speaking English either “very well” or “less than very well” from 2015 through 2019.

Item 1: Quality of English Spoken by People Who Speak Spanish at Home
in the United States, 2015–2019

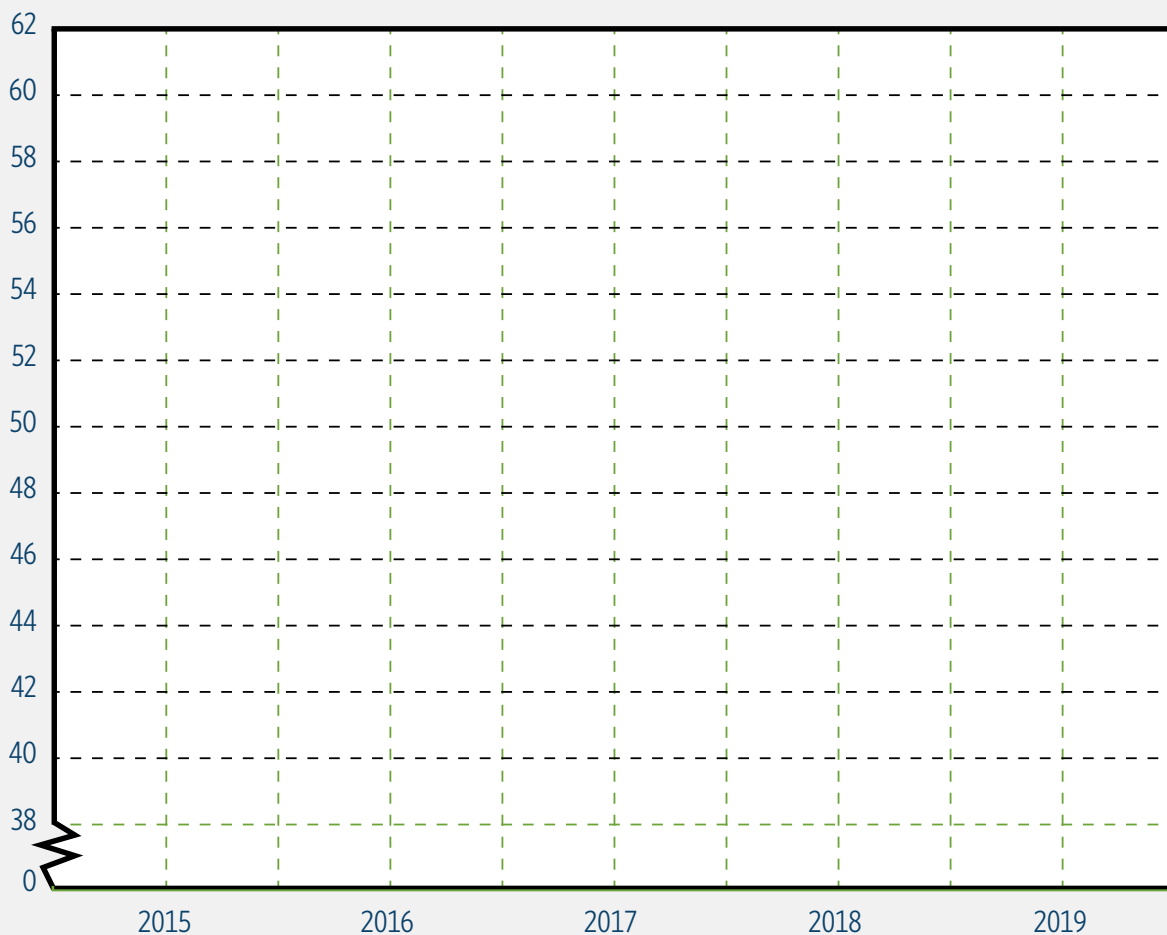
	2015	2016	2017	2018	2019
Percentage who reported speaking English “very well”	59.0	59.0	60.0	60.8	61.4
Percentage who reported speaking English “less than very well”	41.0	41.0	40.0	39.2	38.6

Source: U.S. Census Bureau, 2015, 2016, 2017, 2018, and 2019 American Community Survey 1-Year Estimates, Table B16001

To represent each variable on your scatter plot, you will need to use two different colored pencils. Fill in the boxes of the legend below with each color.

Legend:	Speaks English "Very Well"	Speaks English "Less Than Very Well"
---------	-------------------------------	--

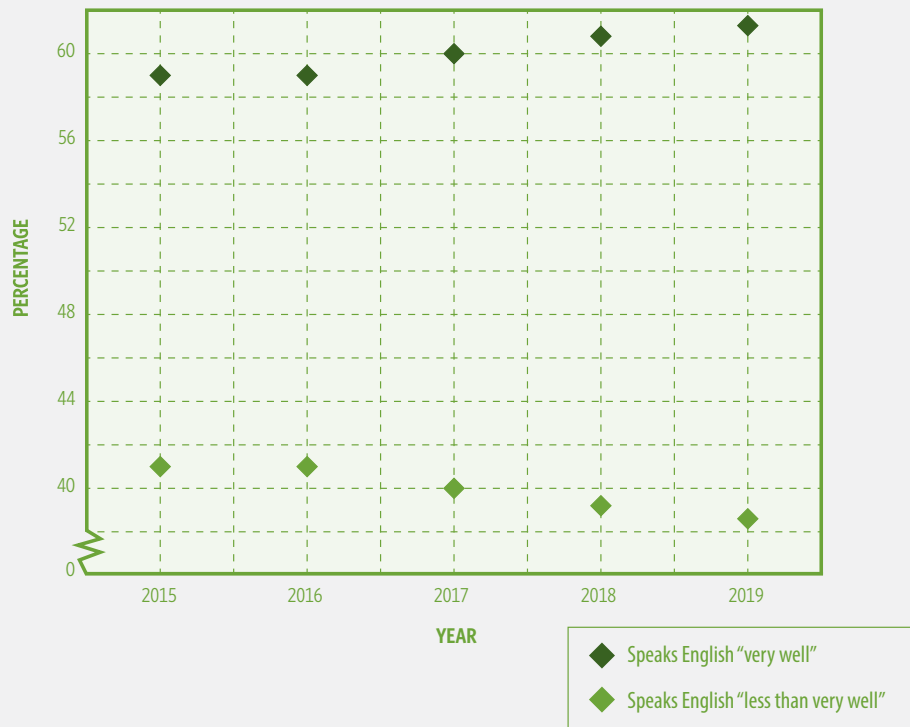
- Now, using the colors that you marked in your legend, place 10 dots on this graph to represent the 10 data points – 5 in each category.



2. Label each of your axes and write a title at the top of the scatter plot.

Student graphs should look like this:

Quality of Spoken English in Primarily Spanish-Speaking Homes in the United States



3. What is the explanatory variable on your scatter plot?

The explanatory variable is the year.

4. What is the response variable for each set of data points?

The response variables are the percentage of Spanish speakers who reported speaking English "very well" and the percentage of Spanish speakers who reported speaking English "less than very well."

5. If data points on a scatter plot form a positive slope, it means that they have a positive association. In these cases, as you look from left to right on the scatter plot, the line appears to move "uphill." This means that as the explanatory variable increases, the response variable tends to increase as well.

- a. Which category of data shows a positive association over time?

The percentage of Spanish speakers who reported speaking English "very well" shows a positive association.

- b. Could you explain why that association may be positive?

Student answers will vary but could include the following: Over time, English-language education has become more accessible in the United States; people tend to become better English speakers as they interact with fluent English speakers; etc.

6. If data points on a scatter plot form a negative slope, it means they have a negative association. In these cases, as you look from left to right on the scatter plot, the line appears to move “downhill.” This means that as the explanatory variable increases, the response variable tends to decrease.

- a. Which category of data shows a negative association over time?

The percentage of Spanish speakers who reported speaking English “less than very well” shows a negative association.

- b. Could you explain why that association may be negative?

Student answers will vary. (See suggested responses for question 5b.)

7. Use a straight edge to draw a best fit line through the data points, or a straight line through the center of the data points, for speaking English “very well.” Do the same to try to connect the data points for speaking English “less than very well.” Which set of data more closely follows a straight line?

Both. The percentage speaking English “very well” and “less than very well” data points more closely follow a straight line.

8. Summarize the association between the percentage of Spanish-speaking residents who reported speaking English “very well” and the years 2015 through 2019:

Student answers will vary but could include: There is a strong, positive, and linear association between the years 2015 through 2019 and the percentage of Spanish-speaking residents who reported speaking English “very well.”

9. Write a question about the scatter plot for a classmate to answer:

Student questions will vary.

10. When your teacher directs you, turn to a classmate near you and ask your question from #9. How well did your classmate answer the question? Did you need to reword the question so that your classmate could understand what you were asking? Did you follow up with another question? Explain.

Student answers will vary.

11. Is there a relationship between the percentage of Spanish speakers who reported speaking English “very well” and the percentage of Spanish speakers who reported speaking English “less than very well”?

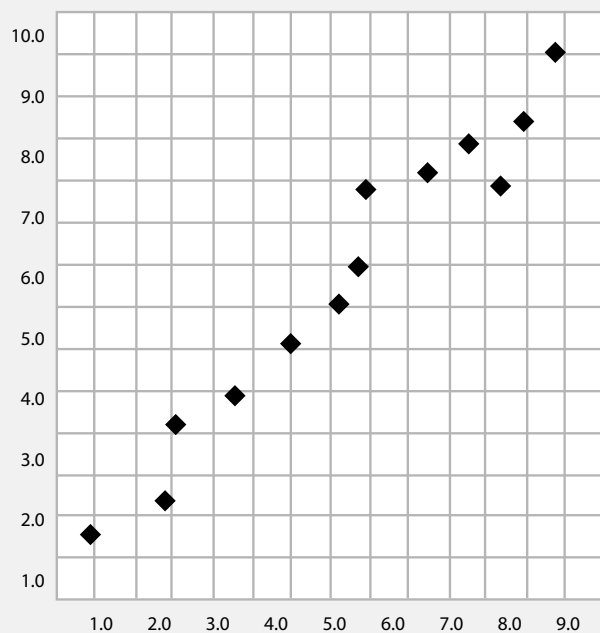
Yes, the two percentages add up to one.

Part 2 – Apply What You’ve Learned

What have you learned about association? Comment on the direction, form, and strength of each association below:

1.

x = hours since rain started to fall
y = centimeters of rain collected in a rain gauge



Direction: **Positive**

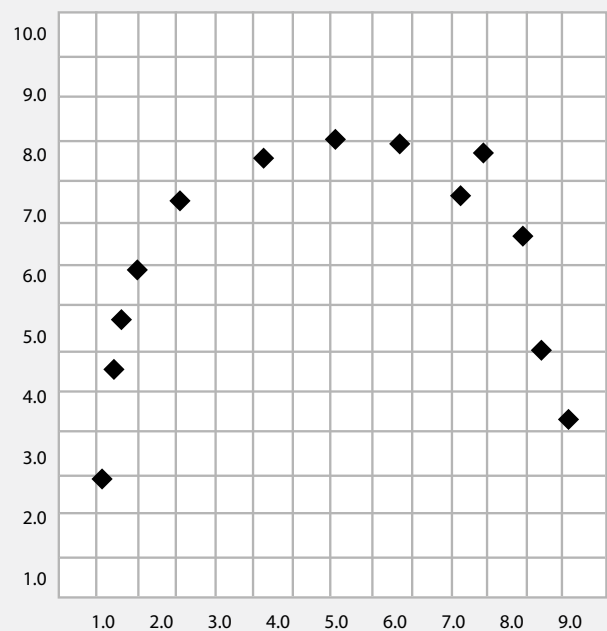
Form: **Linear**

Strength: **Strong**

Explanation: **The longer it rains, the more centimeters of rain will be collected in the gauge. In this example, it seems that the rate of rainfall is somewhat constant.**

2.

x = weeks since a video game was released
y = number of video games sold by one Internet vendor



Direction: **Positive, then negative**

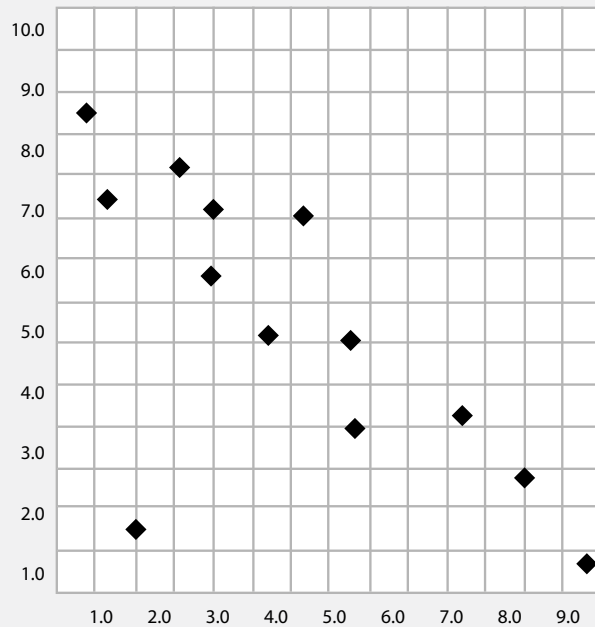
Form: **Curved**

Strength: **Strong**

Explanation: **Student answers will vary but could include the following: The form is parabolic or quadratic, meaning that it is curved. When a video game is first released, it is not very popular because not many people know about it. However, as more people learn how much fun it is, the sales increase. Eventually, everyone who wants the game already owns it, so sales drop off.**

3.

x = number of hours spent practicing recital music
y = number of mistakes made (as recorded by music instructor)



Direction: **Negative**

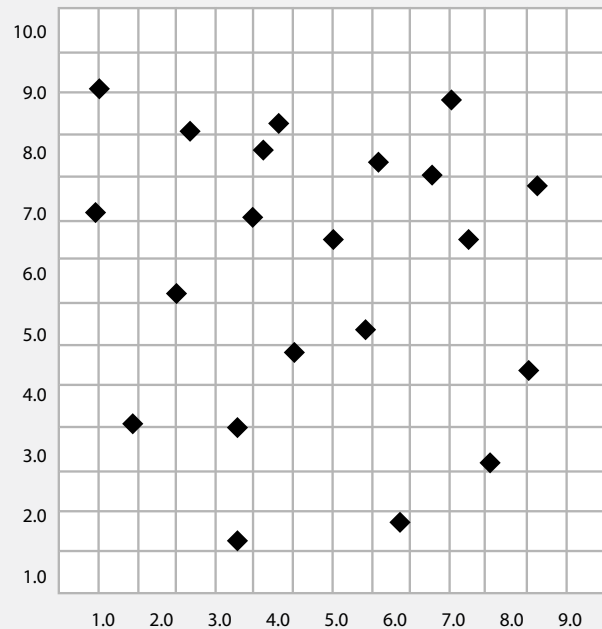
Form: **Linear**

Strength: **Somewhat strong (excluding the outlier)**

Explanation: **As the musician practices over time, he or she makes fewer mistakes (as recorded by the instructor).**

4.

x = number of baseball cards owned by a student
y = number of minutes it takes a student to walk to school



Direction: **None**

Form: **None**

Strength: **Weak**

Explanation: **Direction, form, and strength indicate that the two variables are not associated. The number of baseball cards that a student owns is not related to the number of minutes it takes for a student to walk to school.**

Part 3 – Determine and Explain Negative and Positive Associations

For each pair of variables, note whether you expect the association to be positive or negative, and explain why:

Scenario	Positive or Negative?	Explanation
1. x = child's age (years) y = child's height (centimeters)	Positive	As a child gets older, he or she tends to get taller.
2. x = days a student misses school y = hours spent on makeup work	Positive	The more school days a student misses, the more makeup work will pile up (meaning more hours spent on makeup work).
3. x = ounces of insecticide used y = mosquito population (hundreds per acre)	Negative	As more insecticide is used, the mosquitoes will likely die off, decreasing their population.
4. x = hours a student spends watching TV y = hours a student spends playing video games	Student answers will vary (could be positive or negative)	Explanation for positive association: Students who like watching TV also may like playing video games, so they could spend many hours doing both. Explanation for negative association: If a student is allowed to spend only a certain amount of hours looking at a screen, more time watching TV means less time for video games (and vice versa).
5. x = hours spent studying the spelling of vocabulary words y = number of vocabulary words spelled incorrectly	Negative	As a student spends more time studying vocabulary words, he or she will likely spell fewer words incorrectly.
6. x = hours spent digging a hole y = depth of the hole (feet)	Positive	If someone spends more hours digging a hole, then that hole will probably get deeper.